# **Sepehr Dehdashtian**

# 517-721-0269 | sepehr@msu.edu | Website | Google Scholar | LinkedIn | GitHub

| Education   |                    |  |  |
|---|--------------------|--|--|
| Michigan State University   Ph.D. in Computer Science   | ın. 2022 – Present |  |  |
| <ul> <li>Focus: Responsible AI, Generative AI, Multimodal Models, Computer Vision</li> <li>GPA: 4.0/4.0</li> </ul>      |                    |  |  |
| Sharif University of Technology   M.Sc. in Electrical Engineering Sep.  | . 2018 – Feb. 2021 |  |  |
| <ul> <li>Thesis: Blind Recognition of Channel Codes Using Deep Neural Networks</li> <li>GPA: 3.87/4.0</li> </ul>        |                    |  |  |
| Shahid Chamran University of Ahvaz   B.Sc. in Electrical Engineering Sep.   | 2014 – Aug. 2018   |  |  |
| • <b>GPA</b> : 3.93/4.0 (ranked 1 <sup>st</sup> )   |                    |  |  |
| Publications  |                    |  |  |
| OASIS Uncovers: High-Quality T2I Models, Same Old Stereotypes   | 2025               |  |  |
| Sepehr Dehdashtian, Gautam Sreekumar, Vishnu Naresh Boddeti   |                    |  |  |
| International Conference on Learning Representations (ICLR) 2025 (Spotlight)  |                    |  |  |
| Fairness and Bias Mitigation in Computer Vision: A Survey   | 2024               |  |  |
| Sepehr Dehdashtian*, Ruozhen He*, Yi Li, Guha Balakrishnan, Nuno Vasconcelos, Vicente Ordonez, Vishnu Naresh Boddeti    |                    |  |  |
| IEEE Transaction on Pattern Analysis and Machine Intelligence (TPAMI) (Under Review                                     | N)                 |  |  |
| The Dark Side of Dataset Scaling: Evaluating Racial Classification in Multimodal  | Models 2024        |  |  |
| Abeba Birhane*, Sepehr Dehdashtian*, Vinay Prabhu, Vishnu Naresh Boddeti  |                    |  |  |
| ACM Conference on Fairness, Accountability, and Transparency (FAccT) 2024   |                    |  |  |
| Utility-Fairness Trade-Offs and How to Find Them  | 2024               |  |  |
| Sepehr Dehdashtian, Bashir Sadeghi, Vishnu Naresh Boddeti   |                    |  |  |
| IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2024  |                    |  |  |
| FairerCLIP: Debiasing CLIP's Zero-Shot Predictions using Functions in RKHSs   | 2024               |  |  |
| Sepehr Dehdashtian*, Lan Wang*, Vishnu Naresh Boddeti   |                    |  |  |
| International Conference on Learning Representations (ICLR) 2024  |                    |  |  |
| On characterizing the trade-off in invariant representation learning  | 2022               |  |  |
| Bashir Sadeghi, <u>Sepehr Dehdashtian</u> , Vishnu Naresh Boddeti   |                    |  |  |
| Transactions on Machine Learning Research (TMLR)  |                    |  |  |
| Deep-Learning Based Blind Recognition of Channel Code Parameters over Candi under AWGN and Multi-Path Fading Conditions | idate Sets<br>2021 |  |  |
| Sepehr Dehdashtian, Matin Hashemi, Saber Salehkaleybar  |                    |  |  |
| IEEE Wireless Communications Letters  |                    |  |  |
| Professional Experience   |                    |  |  |
|   |                    |  |  |

• Identified failure modes of deepfake detectors using LLM fine-tuning with RL and Agentic Al.

Dec. 2024 - Present

Research Intern at Reality Defender (Mentor: Dr. Jacob Seidman)

#### Research Assistant at Michigan State University

Jun. 2022 - Present

- Developed algorithms to make computer vision, multimodal, and generative models fair and debiased.
- Published papers in top computer vision and machine learning conferences: ICLR'24, CVPR'24.

## Research Assistant at Sharif University of Technology

Nov. 2018 - Feb. 2021

- Developed deep learning based methods for blind recognition of channel codes.
- Published a research paper in IEEE Wireless Communication Letters.

#### **Student Intern at Radaq Company**

Jun. 2017 - Aug. 2017

2022

- Designed a data logger system for agricultural sensors.
- Self-studied PHP and HTML to develop a website for the data logger system

| <b>Awards &amp; Honors</b> | Aw | ards | & I | Honors |
|----------------------------|----|------|-----|--------|
|----------------------------|----|------|-----|--------|

| • ICLR 20 | 25 Spotlight Paper (Top 5%) | 2025 |
|-----------|-----------------------------|------|
|-----------|-----------------------------|------|

STEAMpower Fellowship
 2024

• TMLR Outstanding Paper Award Runner-Up and TMLR Featured Certification Award 2023

Ranked 2nd GPA the graduating class of 2021 | Sharif University of Technology

• Ranked 5th out of 30,000 | Iranian University Entrance Exam for Master's degree 2018

• Ranked 1st GPA the graduating class of 2018 | Shahid Chamran University of Ahvaz 2018

#### Technical Skills -----

Languages: Python, C++, CUDA, Verilog, VHDL

ML Frameworks: PyTorch, PyTorch-Lightning, TensorFlow, Keras, OpenCV, Scikit-Learn

Mitigating Political Piec in Pro Trained Large Language Models

Others: RevealJS, Git, MATLAB, FPGA

## Projects ----

| • | witigating Political Bias in Pre-Trained Large Language woders       | 2023 |
|---|--|------|
| • | Visually Explaining Fair Representation Learning—A Model Perspective | 2023 |
| • | Video Synopsis using OpenCV in Python                                | 2019 |

#### Services & Activities -----

Mentored Student: Yilin Zheng (Master Student at MSU)
 Jan. 2024 – present

• Scientific-Students Associations of Clean Energy | Deputy Secretary 2016

• Scientific-Students Associations of EE Department | Member 2015